



**Making data science pay:
Mastering the challenges of analytics operations**

Michel Debiche
Guest Analyst, STAC

michel.debiche@STACresearch.com

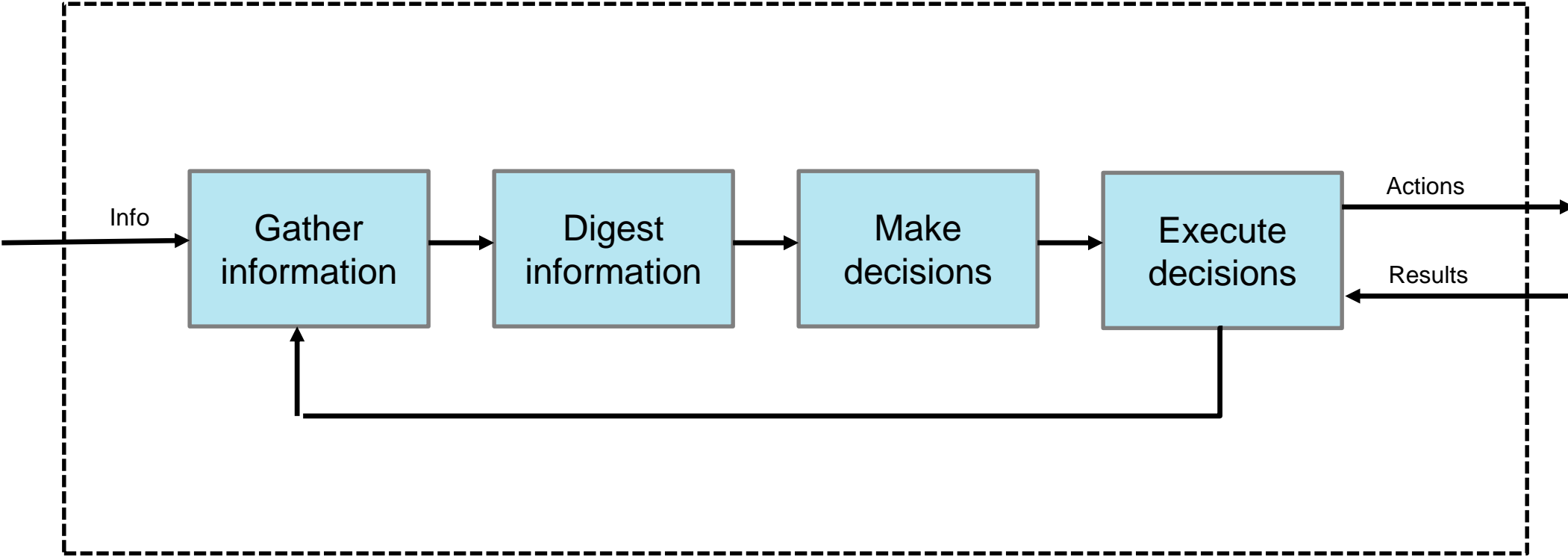
Cognitive Reset Part 1: Window management technology



Cognitive Reset Part 2: Window management context



Investment Process



Pressures

- Scale
 - Volume
 - Variety
 - Density
 - Computational complexity
 - Velocity of innovation
- Cost
- Regulation

Dimensions of scale

- **Scale**

- **Volume**

- **Variety**

- Kinds of data: structured, unstructured, text, binary
 - Data entities: Millions of time series

- **Density**

- Transactions in microseconds
 - Simultaneous transactions on multiple channels

- **Computational complexity**

- NLP, Image processing, AI

- **Velocity of innovation**

- Competitive pressures: New datasets, new models, new technologies
 - Evolving opportunities
 - Feedback loops

Responses

- DevOps
- Data Lake
- Open Source
- Big Data
- Data Science
- AI

Issues

- Model Factories: Hundreds of models with nowhere to go
- Redundant engineering
- Open source interoperation and upgrade nightmares
- Murky, expensive data lakes contributing little value
- Skills mismatches
- User resistance to new technologies
- Data lineage, audit trails

Goals

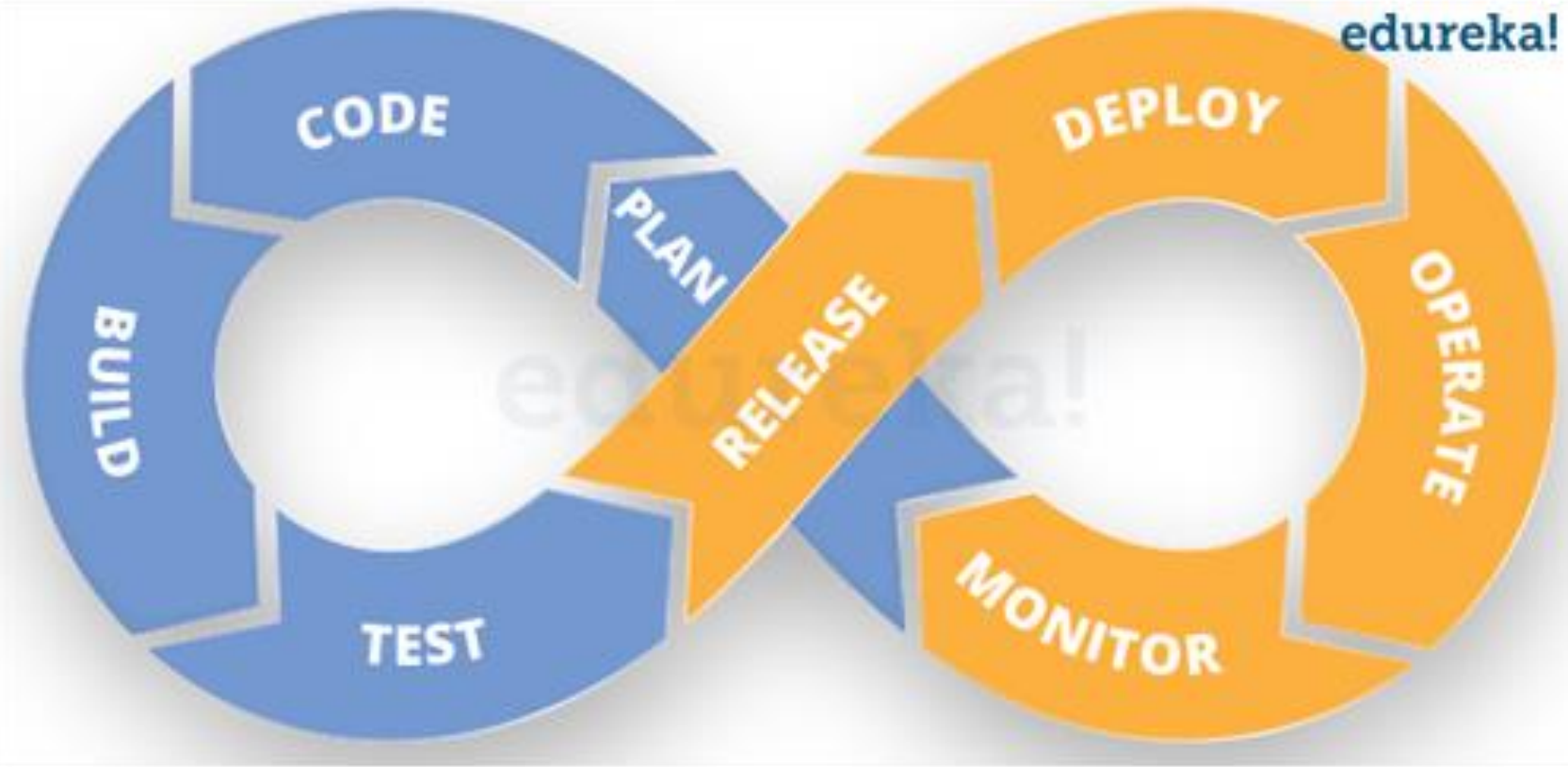
- Maximize returns
- Minimize risk
 - Market risk
 - Model risk
 - Systems risk
 - Data risk
 - Operational risk (people)
- Maximize productivity

Principles

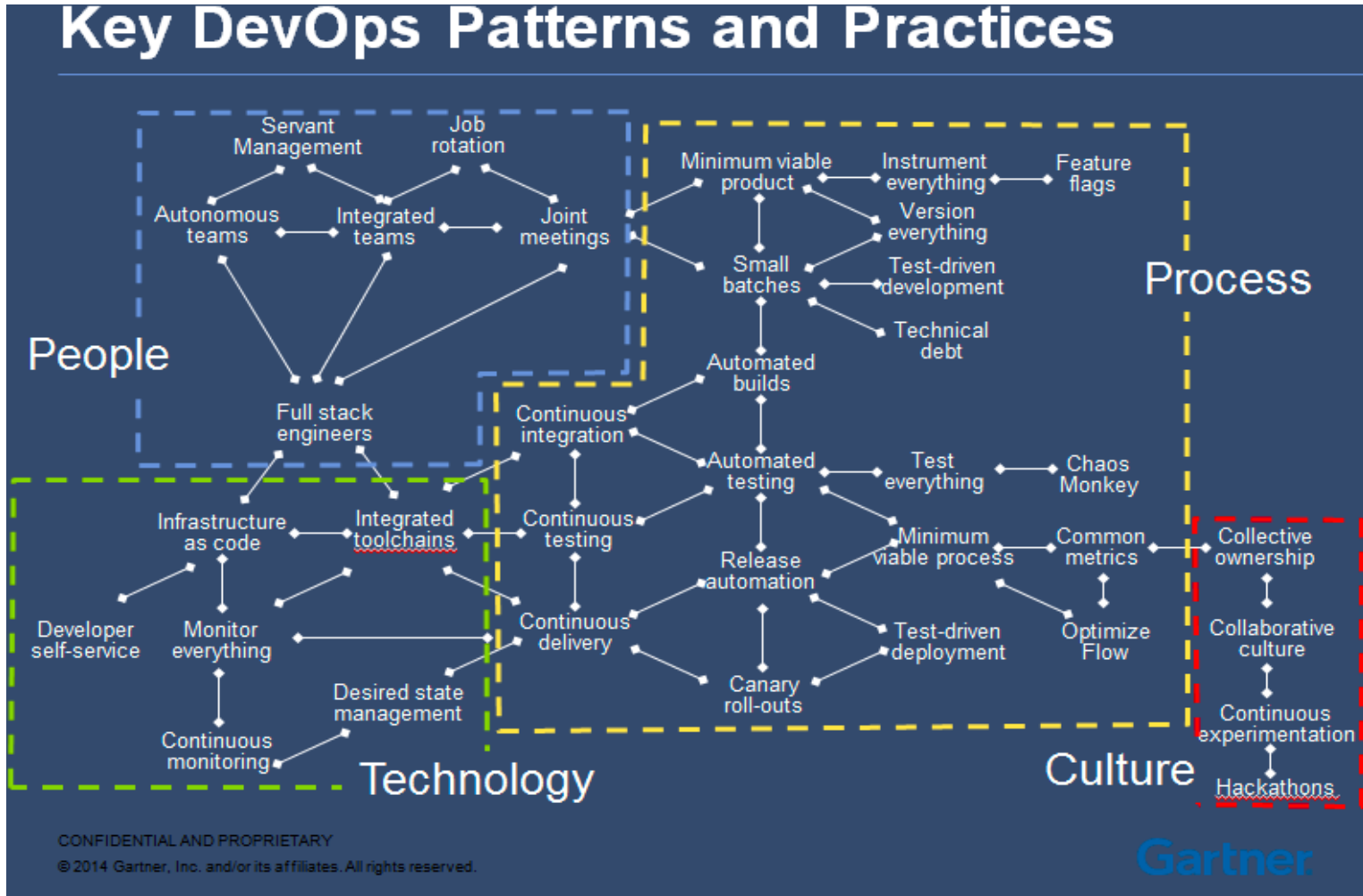
- Optimize use of resources
 - People
 - Time
 - Data
 - Technology
- End-to-end process design
- Agility
- Constant improvement

- Similar challenges and goals
- Eventually came to software engineering as DevOps
- Need to carry paradigm over to full data-to-decision pipeline
- Why is it so hard?

DevOps: Elegant Concept

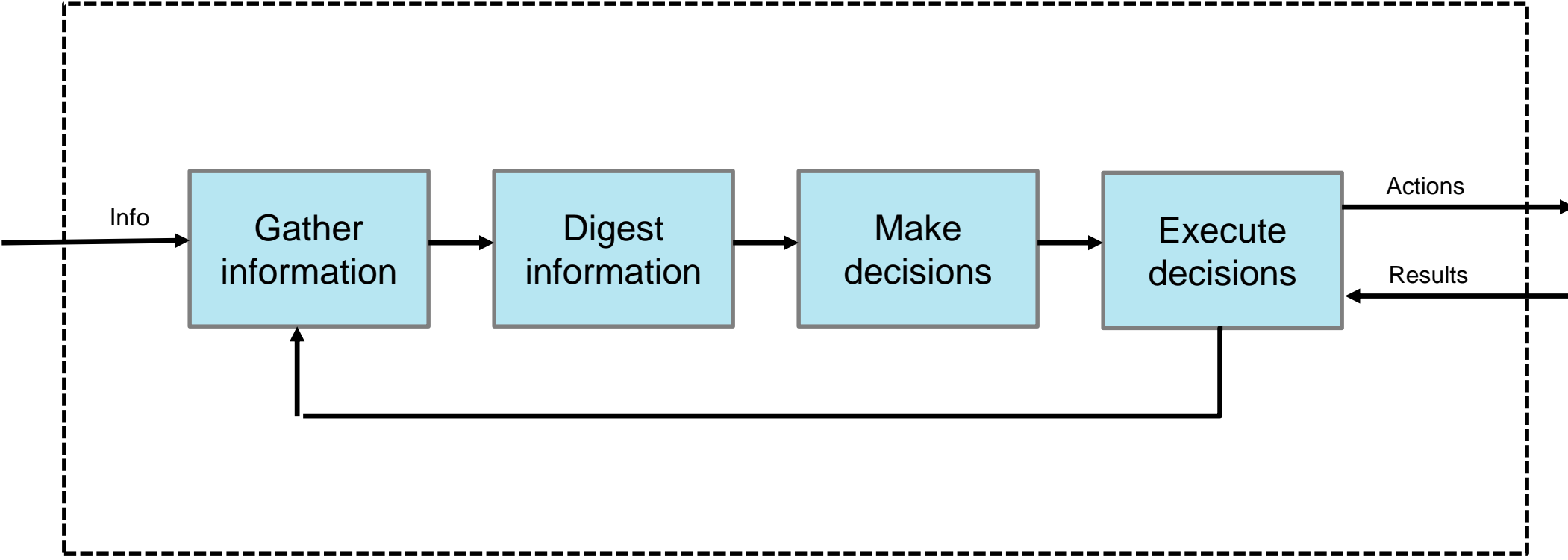


DevOps: More complicated to implement

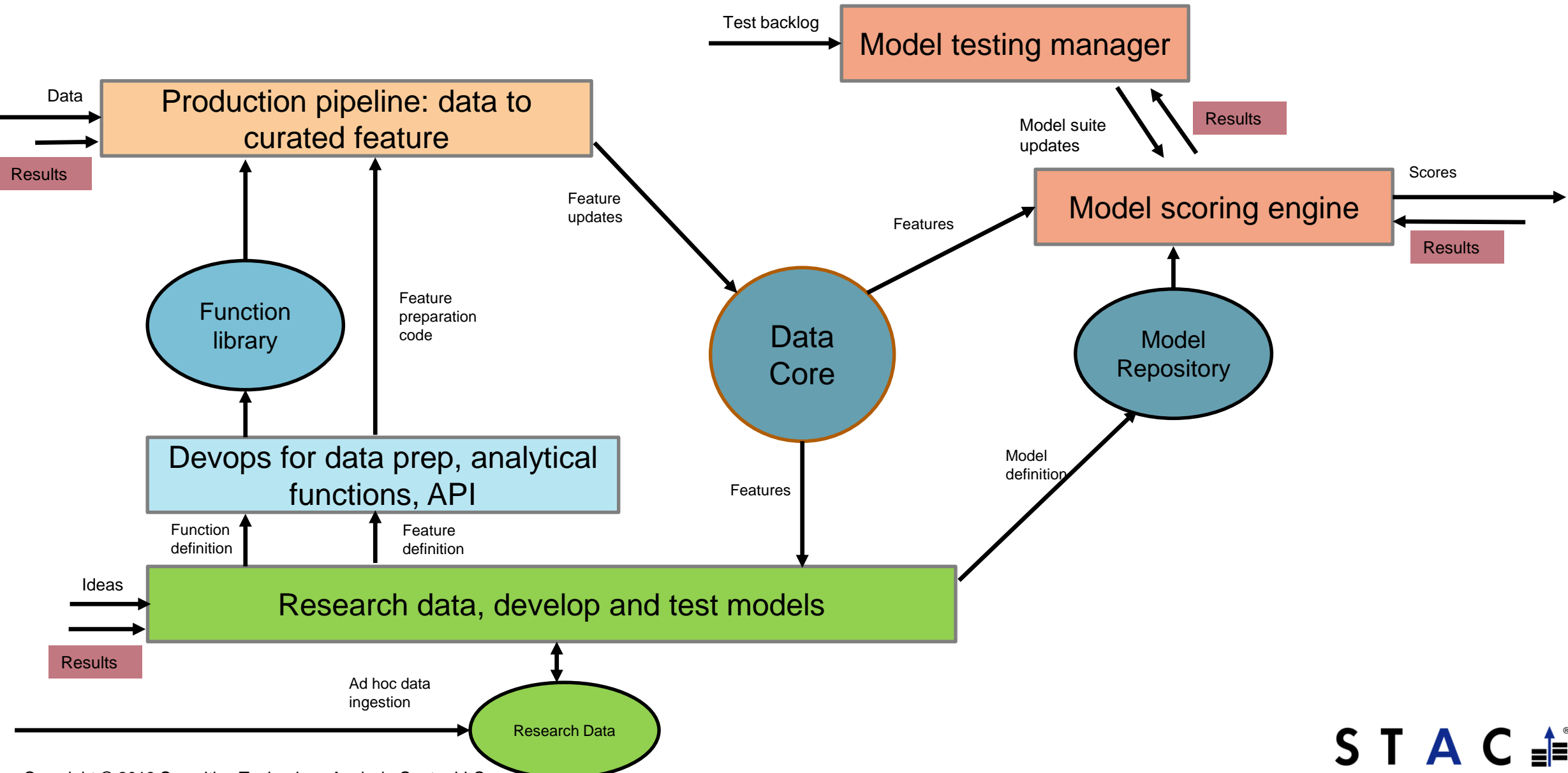


So let's think about QuantOps™

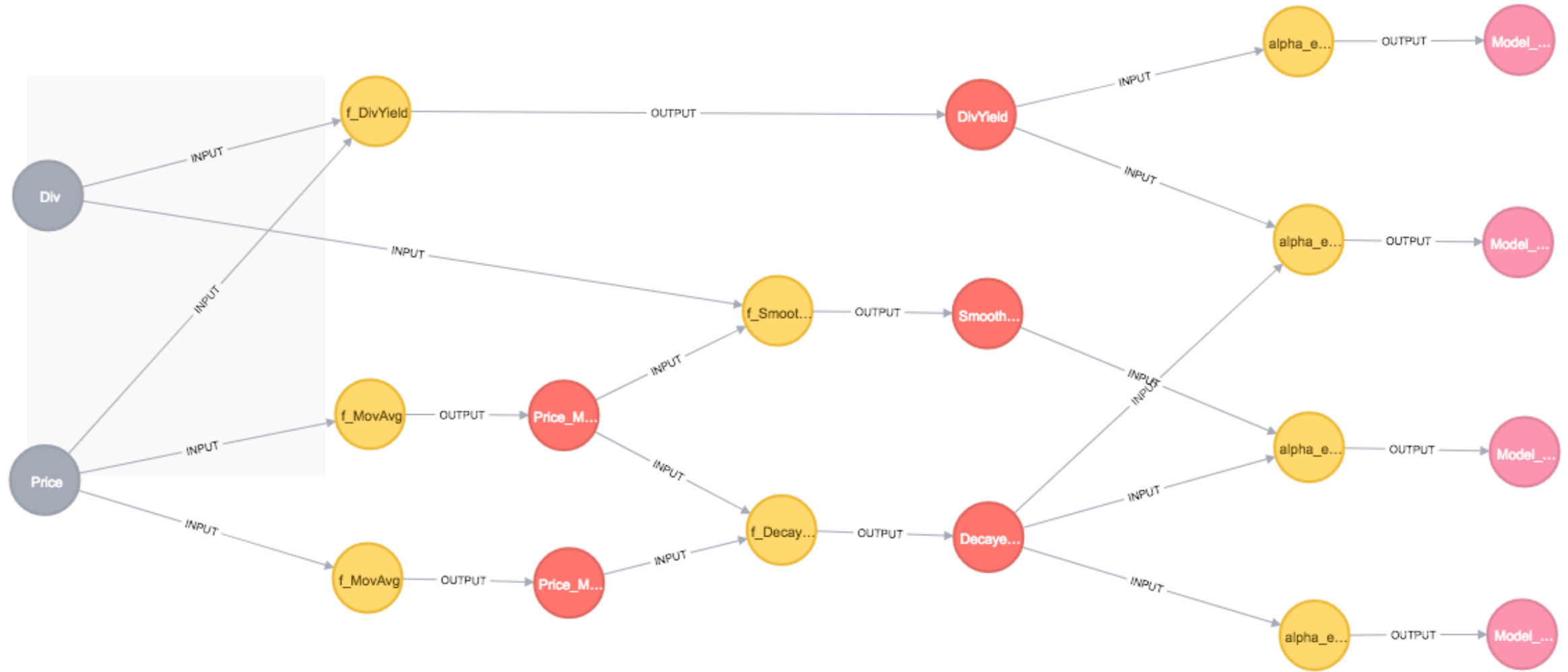
Investment Process



Investment Process, Expanded



A Unifying Paradigm: QuantOps as a DAG



A Unifying Paradigm: QuantOps as a DAG

- Standardize the connections
- Carefully define the data APIs
- Then all the technology is pluggable
- Makes it possible to efficiently address:
 - Orchestration
 - Data lineage
 - Monitoring
 - Audit trails
 - Automated code generation and testing

Where does STAC fit in?

- Implementing analytics ops is a big commitment with big payoffs
- Biggest challenge: effective communication, change management
- Design needs to be process-oriented and based on user needs
- Technology needs to respond to process requirements, not vice versa
- Emerging STAC roles:
 - Facilitate dialogue & training on analytics ops challenges & best practices
 - Accelerate technology selection based on community-source standards driven by process-oriented model of the investment process
- Let us know if you want to be involved!